TOWARDS A RECONCILIATION BETWEEN THE ASSOCIATIONIST AND RADICAL BEHAVIORIST TRADITIONS IN THE EXPERIMENTAL ANALYSIS OF BEHAVIOR.[1]

Ullin T. Place, *University of Wales, Bangor.*

Abstract

It is an implication of the Law of Non-Contradiction that two incompatible descriptions of the same class of phenomena cannot both be true. This suggests that the future for radical behaviorism must lie in achieving a reconciliation with other disciplines and approaches studying the same or closely related phenomena. The approach known as "associative learning theory" shares a common data basis with radical behaviorism in the area of the experimental analysis of animal behavior. It is separated from radical behaviorism by a different view of the nature of what is learned. According to the radical behaviorist, under certain antecedent conditions (discriminative stimulus + establishing condition) an organism learns to emit a response. According to associative learning theory what is learned is an association between a pair of consecutive stimulus events. When presented with the first member of the pair, the organism learns to "predict" or "expect" the second member of the pair.

Until recently, the principal application of this principle was Rescorla and Wagner's (1972) analysis of Pavlovian (respondent) conditioning. More recently, Adams and Dickinson's (1981) reinforcer-devaluation experiment has led associationists to pay more attention to instrumental (operant) learning. It has also opened up an interesting divergence of views between Dickinson (1988; Heyes and Dickinson, 1990; Dickinson & Balleine, forthcoming) who takes it as evidence of a discontinuity between respondent conditioning, which he interprets in terms of the establishment of mechanical associations, and operant learning which he interprets in terms of the 'beliefs' and 'desires' of philosophical action theory, and Rescorla (1991) who uses it as evidence for an interpretation of operant learning based on the same principles of stimulus-stimulus association invoked by Rescorla and Wagner to account for respondent conditioning.

Standing in the way of a reconciliation between radical behaviorism and associative learning theory are the misgivings of the former about the use made by the latter of 'mentalistic' concepts, such as 'expect,' 'anticipate,' and 'predict.' These misgivings may be allayed if attention is paid to the results of applying to such concepts the technique, known as 'conceptual analysis,' developed by Wittgenstein (1953; 1958) and the philosophers of the Oxford 'ordinary language' school. A recent application of this technique to the linguistic phenomenon known variously as 'intentionality' or 'intensionality' shows that it consists of two distinct varieties of 'referential anomaly' which 'infect' the grammatical objects of certain verbs. In one case, the grammatical object is used to indicate a range of possible events any one of which, if it were to occur, would constitute a manifestation or satisfaction of a *disposition*. In the other case, the grammatical object functions as a *quotation* of what the agent either has said or might be expected to say or have said. Referential anomaly of the dispositional kind is both unavoidable and benign, but the use of quotations to characterize behavioral dispositions is acceptable for scientific purposes only in those cases where the behavior in question is in fact subject to linguistic control.

Since the grammatical object of the verbs 'know,' 'believe' and 'think,' as they occur in belief/desire explanations, takes the form of an embedded indicative sentence in *oratio obliqua* or indirect reported speech, Dickinson's explanation of instrumental/operant learning in animals involves the scientifically unacceptable metaphor of linguistic initiation and control. Rescorla's theory, on the other hand, requires nothing more than that the organism learn to 'expect' or 'anticipate' an event (the *outcome*), given the combination of an antecedent discriminative *stimulus* and the stimulus constituted by the incipient emission of the *response* which it evokes. In this case the anomaly of reference in the noun phrase which occurs as the grammatical object of the verb reflects its use as a device for indicating a range of *possible* outcomes any one of which, if it occurred, would fulfill and confirm the expectation which it specifies.

Introductory homily

One of the more deplorable features of the intellectual climate as the twentieth century draws to its close, is a loss of confidence amongst scientists and humanists alike in the human ability to distinguish the true from the false. In the case of the humanists, this loss of confidence is associated with a deep-rooted subjectivism which, when combined with the commendable moral principle of toleration for diverse ideological and religious perspectives leads to a thoroughgoing relativism in which what is true is simply what one individual

---

or group holds to be true. There is no fact of the matter. In the case of the scientist and the philosopher of science, the loss of confidence is paradoxically the result of the success of science in discovering what appears at first sight to be remarkable new truths about the world. One need only consider the changes that have occurred in scientific conception of the origin of the universe within the lifetime of even the youngest member of this audience to be led to wonder what further changes in conception will have come about a hundred years from now, assuming that scientific knowledge continues to accumulate at an expanding rate, as it has done for the past three and a half centuries. How can we be sure that the scientific propositions and theories which we now take to be matters of established fact will not turn out to be just as much a mistake as some of the myths and errors of our scientific past?

The science of behavior is perhaps an area where such a radical revolution in our scientific understanding is rather less likely than in fields such as astronomy and sub-atomic physics. At the same time a healthy concern with objectivity in scientific observation which has always been the hallmark of the behaviorist should protect us from the excesses of relativism and subjectivist scepticism which infects the more mentalistically orientated social sciences, as well as much literary and humanistic thinking.

Sadly, some of the things that Skinner wrote in his book *Verbal Behavior* (Skinner 1957) show that he too was deeply sceptical about the possibility of arriving at any kind of truth through the methods of science or indeed, of finding any kind of place for truth within a science of verbal behavior. "Have I told him (i.e. the reader) the truth?" he asks in the first of two epilogues to the book "Who can say?" he continues "A science of verbal behavior makes no provision for truth and certainty (but we cannot even be certain of the truth of that)" (Skinner 1957 p. 456)

Part of the explanation both of Skinner's tendency to think of truth as a matter of concern only to the more pedantic variety of philosopher and of his inability to see a place for the theory of truth as part of a science of behavior is that in *Verbal Behavior* he looks at the language almost exclusively from the standpoint of the speaker and the factors controlling the speaker's utterance. It was not until shortly before his death (Skinner 1989) that he addressed the problem of the response of the listener which is where the problem of truth arises. It arises, not as a mere academic pursuit of concern only to pedants; but as a matter of life and death to the listener who needs to know which of the contingencies specified by the verbal behavior of another speaker or writer, she can rely on in selecting a behavioral strategy appropriate to the situation with which she is confronted. In his essay on 'The behavior of the listener' Skinner has some perceptive things to say about the fact that, as listeners, we tend to trust those persons and types of utterance that have "less often misled us" in the past, and about the value to both parties of the listener's agreeing with what a speaker says. What he does not address is the question of how the listener learns to protect him or herself from the often disastrous consequences of acting on the lies and other forms of misinformation supplied by others. It is true, as Skinner implies, that to some extent we learn by bitter experience of the consequences to trust what some people say and to mistrust others. But that is a long way from being the whole story.

What we need and what we get is advance warning that the information that is being supplied to us is not what it purports to be, an accurate specification of contingencies operating in parts of the environment otherwise inaccessible to us. For this purpose the only principle on which we can ultimately rely is the principle of the *indivisibility of truth* or 'holism,' as it is sometimes called. This is the principle according to which every true proposition must be consistent with every other true proposition. It is a straightforward consequence of the law of non-contradiction whereby, if *p* is true, *not p* must be false and *vice versa*. It follows from this law that if *q* entails *not p*, *p* and *q* cannot both be true. Either one is true and the other false or both are false. And from this it follows that, in building up a stock of beliefs about the world on which to base one's action, one should be made uncomfortable when one detects a contradiction or 'cognitive dissonance,' as Leon Festinger (1957) calls it, within one's existing belief system and endeavour to ensure that any such contradiction is ironed out, before the relevant beliefs are accepted as reliably true. The effect of this endeavour should be to ensure that by and large an individual's beliefs will constitute a coherent system and, provided most of constituent beliefs are true, will thereby constitute a body of knowledge whose reliability will be confirmed by its overall utility as a guide to action (the pragmatic principle) and its conformity to the opinions of others (Wittgenstein's, 1953, §242, "agreement in judgments").

Given such a coherent body of beliefs whose overall correspondence with reality is guaranteed by its consistent reliability as a guide to action, the individual has a standard against which to evaluate any new piece of putative information presented to it by another speaker. If there is no obvious dissonance or

contradiction between the new item and the existing stock, it can be allowed to go through on the nod. Only when a contradiction or dissonance is detected between the new item and the existing stock will alarm bells ring and all the armoury of logical argument be brought to bear in order either to justify the new item's rejection or find some way of resolving the contradiction and incorporating the new item into the system.

How easy or difficult it is to resolve the dissonance and thus incorporate a new belief into one's system will depend on how central to the system as a whole are the existing beliefs which are contradicted by the new candidate belief, and which have to be either modified or abandoned, if the new candidate belief is to be taken on board. Beliefs that are central to the system in this sense are ones which have or appear to the believer to have tight logical links with a large number of other important beliefs. So much so that it is impossible, as things stand, to abandon the belief in question without also abandoning the other beliefs with which it is logically tied. Not surprisingly most people find the effort and ingenuity required to undertake such a radical reconstruction of their belief system beyond them and as a result are inclined to avoid interacting verbally with those who maintain and assert opinions which stand in radical cognitive dissonance with beliefs which are central to their own system. It is this, I suggest, which explains the phenomenon to which Skinner draws attention in 'The behavior of the listener' when he remarks that in conversation:

> Speaking is reinforced when the listener tends to say more or less what the speaker says, and listening is reinforced when the speaker says more or less what the listener tends to say. Conversing is not reinforced by the consequences (external to the conversation) we have been considering but by agreement (Skinner *op. cit.* p. 94).

This tendency to avoid cognitive dissonance, to associate only with those whose views one shares, is not restricted to ordinary social contacts. It affects the behavior of scientists, especially those in behavioral sciences where the scientist's central belief system has direct implications both for one's personal treatment of others and for the political organisation of society. It has deeply affected our own movement in recent years. Faced with the capitulation of every discipline with which it has natural affinities to the so-called "cognitive revolution" which takes the repudiation of the central tenets of behaviorism as axiomatic, behavior analysis has become almost totally isolated not only from other approaches within its parent discipline, but from every other related discipline as well. As I put it some years ago in an article in what has since become *The Analysis of Verbal Behavior*, behavior analysis

> has been consigned to a kind of academic ghetto - cut off by mutual suspicion and incomprehension not only from other approaches within psychology, but from virtually every other adjacent discipline from philosophy, linguistics and sociology on the one hand to ethology and the neuro-sciences on the other. (Place, 1985, p. 38)

This partly self-imposed isolation, implying as it does, that we have monopoly of truth and that everyone else is wrong, cannot be healthy. The principle of the indivisibility of truth demands that sooner or later behavior analysis is going to have to re-establish old links and forge new ones with all these disciplines and take its rightful place as the essential link between the biological sciences on the one hand and the social sciences and humanities on the other. There are some signs, I believe, that the time to start building and rebuilding those bridges is already upon us. The connectionist movement in artificial intelligence and the selectionist movement in biology in general and neuroscience in particular (Place 1992a; 1992b) are undermining the hold of the serial-digital-computer as a model for the functioning of the brain, and creating a climate in which many of the barriers to communication between behavior analyst and practitioners of other disciplines and approaches within the area of the brain and behavioral sciences are coming down.

Associative learning theory in relation to Behavior Analysis
It is in this spirit and with this objective in mind that I am asking you to explore with me the possibility of achieving a reconciliation between behavior analysis and the approach to the study of the behavior of non-human pre-linguistic organisms - animals - known as associative learning theory.

Although, given the choice, associative learning theorists would probably describe themselves as cognitivists rather than behaviorists, associative learning theory is essentially an offshoot from the same behaviorist root as behavior analysis. It shares with behavior analysis -

(1)   an interest in what is *common* to the learned behavior of different animal species rather than the species-specific innate behavior which preoccupies the ethologists,

(2)   a common data base of experimental studies of animal behavior,

(3)   a common experimental methodology, including in particular the use of the Skinner box and the associated schedules of operant reinforcement,

(4)   the use of many of the same concepts, 'stimulus' and 'response,' together with concepts derived from Pavlov (1927), such as 'reinforcement,'[2] 'conditioned' and 'unconditioned stimulus/response', 'generalisation' and 'discrimination' as well as terms which are readily translated. Associative learning theorists, for example, use Hilgard and Marquis's (1940) distinction between 'classical Pavlovian' and 'instrumental conditioning', where behavior analysts follow Skinner (1938) in speaking of 'respondent conditioning' and 'operant conditioning'.

Associative learning theory differs from behavior analysis in the following respects

(1)   Associative learning theorists are interested in animal behavior primarily for the sake of the light it throws on the brain mechanisms involved, not, as is the behavior analyst, for the sake of the light it throws on the way human behavior is organised at the molar level.

(2)   Unlike behavior analysts, associative learning theorists pay virtually no attention to data from the experimental analysis of human behavior.

(3)   The only practical applications which the associative learning theorist envisages for his or her work is in fields such as psycho-pharmacology and neurology; possible implications for applied behavior analysis, whether with animal or human clients, play no part.

(4)   Associative learning theorists regularly talk in what most behavior analysts regard as an unacceptably 'mentalistic' way about what an animal is 'attending to', 'knows about', 'believes about', 'expects', 'predicts', 'remembers' and 'recognises', and, in the jargon of the cognitive psychologists, about 'retrieving items of information from the memory store'.

Another difference which is more a matter of emphasis is that, traditionally, the emphasis in associative learning theory has been on Pavlovian or as the behavior analyst would say 'respondent' conditioning, rather than on instrumental/operant learning and behavior; whereas in behavior analysis the reverse is the case.

The Rescorla-Wagner model

The effect of these differences and the reason why it is important for the behavior analyst to pay attention to the results of research within the associative learning paradigm is that, on the whole, the experimental analysis of animal behavior within the associationist tradition is conducted at a more detailed and finegrained level than is that within the behavior analytic framework. Two pieces of work stand out, one theoretical, the other experimental. The theoretical paper which is generally accepted as providing the foundation for associative learning theory as a distinctive theoretical standpoint is Rescorla and Wagner's (1972) an analysis of classical/Pavlovian/respondent conditioning. What is distinctive and important about Rescorla and Wagner is that it challenges the notion which is deeply ingrained in the thinking of Hullian neo-behaviorists and radical behaviorists alike that what the animal learns in the classical conditioning experiment is to respond in an old way to a new stimulus.

In Pavlovian conditioning, according to Rescorla and Wagner, the animal does not learn to respond in the presence of the CS in the manner in which it has previously responded in the presence of the US; what it learns is to 'predict' (I would say 'anticipate' or 'expect') the appearance of the US, when presented with the CS. It learns to do this in accordance with the Law of Association by Contiguity which has a pedigree which can be traced back through the British Associationists of the 18th and 19th centuries back to Aristotle and Plato. In the form in which it is used by Rescorla and Wagner, it may be stated as follows:

The Law of Association by Contiguity

---

[2] In associative learning theory the term 'reinforcement' comes to mean something rather different from what it means in behavior analysis. In behavior analysis 'a reinforcer' is an event which strengthens an organism's propensity to emit a response. It corresponds to what is said in ordinary language to 'encourage' the organism to repeat what has just been done. In associative learning theory, 'a reinforcer' is the event whose occurrence the conditioned stimulus 'predicts.' It corresponds to what is said in ordinary language to 'confirm' that expectation.

If two stimulus event types S_1 and S_2 impinge on an organism's receptors in close temporal proximity such that the onset of S_1 is regularly followed after a short and constant interval by the onset of S_2, the organism will learn to *expect* S_2 whenever S_1 is presented.

It must be assumed that associative connections between successive stimulus events are being formed in accordance with the Law of Association by Contiguity, whenever an event of one type is regularly experienced following an event of another type. Under normal circumstances the existence of these associations is manifested, not so much in what the organism *does*, as in what it fails to do, in what it ignores, rather than what it attends to. What attracts attention, apart from the motivationally significant, is the *unexpected* - the case where S_1 is followed, not by S_2, but by a stimulus of another kind S_3 or by nothing happening at all.

Viewed from this perspective, what Pavlov (1927) succeeded in doing was to devise a technique for studying the formation of these input to input associations experimentally. He did this by adventitiously selecting as his S_2 or "unconditioned stimulus" (US) a stimulus event which is of motivational significance to the organism, such as the injection of food (attractive) or acid (aversive) into the mouth, and arranging to record the resultant autonomic response (in this case salivation). This enabled him to use the anticipatory autonomic response (the CR) which develops under these conditions in response to S_1 (the CS) as a way of measuring the expectation which results from the repeated occurrence of S_1 (the CS) immediately followed by S_2 (the US) order on previous trials.

This was not how Pavlov and his followers, including Skinner, construed the matter. He and they made the mistake of supposing that what is important in Pavlovian conditioning is the stimulus control that S_1 (the CS) acquires over the autonomic response originally elicited by S_2 (the US), when, if Rescorla and Wagner are right, the conditioned autonomic response is nothing more than a convenient device for demonstrating and studying the development of an underlying expectation or input to input association.

Evidence supporting the Rescorla-Wagner model

The adoption of the classical Law of Association by Contiguity in the form of the Rescorla-Wagner model as the basis for a programme of experimental research in animal learning has led to a re-assessment of many otherwise puzzling phenomena of learning, both ones whose existence has been known for many years and others which have been discovered more recently. These phenomena fall into two groups. In the first group are phenomena such as 'masking' (Freeman and Thomas, 1967), 'overshadowing' (Pavlov 1927) and 'blocking' (Chase 1968). In these cases one element of a stimulus compound will, depending on the particular training conditions, mask, overshadow or block the stimulus control otherwise exercised by another element of the same stimulus compound. These effects are produced in part by the greater intensity or 'salience' of the dominant stimulus element. But there is also evidence - which only makes sense within the Rescorla-Wagner framework - that an element of a stimulus compound will be masked, overshadowed or blocked if it is a less reliable predictor of subsequent events than the dominant stimulus element, or, in the case of two successive stimuli, if the second stimulus adds nothing to the prediction that is possible on the basis of the first.

In this first group of learning phenomena it is the bias in favor of more reliable prediction/expectation which supports the Rescorla-Wagner interpretation. What unites the second group is the prediction from the Law of Association by Contiguity that associative links will be formed between successive stimuli, regardless of whether or not a response is evoked from the organism They all involve presentations of stimulus events or pairs of such events in which no response is required from the subject at the time; but where the course of subsequent instrumental learning shows that associative links have nevertheless been formed. The learning phenomena in this second group include sensory pre-conditioning (Brogden 1939), latent learning (Tolman and Honzik 1930), autoshaping (Brown and Jenkins 1968) and recognition learning (e.g., MacPhail and Reilly 1989).

The reinforcer-devaluation experiment

After Rescorla-Wagner the second major milestone in the evolution of associative learning theory was an experiment reported in the *Quarterly Journal of Experimental Psychology* in 1981 by Christopher Adams and Anthony Dickinson from the Psychological Laboratory, University of Cambridge. In this experiment the delivery of a particular foodstuff is made contingent on a response (lever-pressing in the rat), and is then devalued outside the experimental chamber by pairing it with gastric illness, induced by an injection of

Lithium Chloride. The result of this procedure is that the propensity to emit the response in the stimulus conditions under which it has been previously reinforced by the delivery of the devalued foodstuff is effectively abolished. When compared with the case in which the reinforcer has not been devalued, extinction of the response when no reinforcer is supplied proceeds much more rapidly and from a much lower initial response rate. Moreover, the effect of restoring the contingency after the extinction phase is to reduce the rate of responding still further, in contrast to the rapid restoration of the previous high level of responding when the reinforcer has not been devalued.

This result is difficult to accommodate within any account which proposes that what is learned as a consequence of instrumental/operant reinforcement is the propensity to respond in a particular way in the presence of a particular set of stimulus conditions. To that extent it supports the claim of the associative learning theorist that what is learned is something 'about' the sequence of stimulus events.

The experiment has had two consequences for associative learning theory:

(1)   It has shifted the focus of interest from an original pre-occupation with classical/respondent conditioning towards an interest in instrumental/operant learning, not, as had previously been the case, simply as a convenient method of demonstrating the operation of classical associative processes, but as a phenomenon of interest in its own right.

(2)   It has opened up a difference of opinion between Tony Dickinson himself (Dickinson 1988; Heyes and Dickinson 1990; Dickinson and Balleine, forthcoming) who takes it as evidence of a radical discontinuity between classical/respondent conditioning on the one hand, and instrumental/operant conditioning on the other, and Bob Rescorla (1991) who, in his 1990 Bartlett Memorial Lecture, takes it as evidence that the principles enunciated by Rescorla and Wagner (1972) in relation to classical/respondent conditioning can be extended to cover instrumental/operant learning.

According to Dickinson the reinforcer-devaluation experiment shows that, unlike classical/respondent conditioning which he thinks can be explained in terms of the formation of mechanical associative links within the brain, instrumental/operant learning involves what he calls "intentionality." In other words, it can only be explained by reference to the beliefs and desires of common sense psychology and philosophical action theory. On this view what the organism acquires in instrumental/operant learning is the "*belief*" that behaving in a certain way will have certain consequences. If those consequences are attractive, it will emit the behavior in question, if they are repulsive it will not. According to the alternative view expounded by Rescorla what is learned is what we may call '*a contingency expectation*'. This is a three component stimulus-stimulus expectation formed in accordance with the same principles of associative learning described by Rescorla-Wagner in the case of classical/respondent conditioning. As a consequence of the repeated pairing of the two, an antecedent stimulus evokes an expectation of the feedback stimulation generated by the response on the part of the organism which has previously been emitted in the presence of that antecedent. By the same token, when the feedback stimulation generated by the incipient emission of the response arrives, it will evoke an expectation of the kind of stimulus event which, in the past, has been the regular outcome or consequence of the emission of that response. Rescorla' theory may be formulated as a deduction from the Law of Association by Contiguity in the form of

> *The Principle of Contingency Expectation*
> Given that the combination of an Antecedent condition with the sensory feedback from
> the Behavior emitted under that condition (S_i) has been regularly and closely followed by
> a certain Consequence (S_2), a network, whether artificial or natural, will learn to expect the
> Consequence (S_2), given the Antecedent condition combined with the sensory feedback
> from an incipient evocation of the Behavior regularly emitted under that condition (S_i).

There is doubtless a temptation for the behavior analyst to react to this dispute between Dickinson and Rescorla over the significance of the reinforcer-devaluation experiment as a private quarrel between associative learning theorists which is of no concern to them. But this is not, I submit a tenable position.

The challenge to the behaviorist's repudiation of mentalism
In the light of the reinforcer devaluation experiment it can no longer be maintained that what is learned in an instrumental/operant learning situation is the propensity to emit a member of a particular response class in the presence of a member of a particular stimulus class. Evidently, what is learned in this type of situation is something *about* the contingency, about the consequence or outcome of behaving in a certain way in the

presence of that kind of stimulus. The behavior actually emitted on a specific occasion will vary, depending on whether the consequences the organism has learned to *anticipate* or *expect* in these circumstances are currently attractive (reinforcing) or aversive. If the anticipated consequences are attractive/reinforcing, the organism will emit the behavior that has produced those consequences in the past. If, on the other hand, the consequences have become aversive by virtue of their association with an aversive event such as illness, any such propensity will be inhibited.

There is, as far as I can see, no way that you can formulate these relationships without

(a)   construing what is learned as some kind of 'expectation' or 'belief about' the consequences of behaving in a certain way, and

(b)   construing the change that occurs whereby a consequence which was previously attractive (reinforcing) becomes aversive as a change in the organism's affective attitude to the consequence in question from liking to disliking.

Alternative responses to the dilemma of mentalism, and Skinner's "gut intuition"

It is no use saying that verbs like *believing, expecting, anticipating, liking,* and *disliking* are all equally mentalistic and therefore, should not be allowed to pass the lips of any self-respecting behaviorist. For if there are phenomena such as this which cannot be described without using some kind of mentalistic language, we are left with only two alternatives. One alternative is to give up the attempt to develop a language for the description and explanation of behavior which goes no further than can be strictly justified by the objective behavioral evidence, and return, with the cognitivists, to the unbridled mentalism which flourished before Lloyd Morgan (1894) promulgated his famous "canon" more than a century ago. That would be to abandon everything that behaviorism stands for and for that reason is not a course which would recommend itself to anyone here. The only other alternative, so it seems to me, is to ask ourselves what it is about mentalism which justifies what the philosopher Dan Dennett (1978) describes as Skinner's (and it is not only Skinner's)

strong gut intuition that the *traditional* way of talking about and explaining human behavior - in 'mentalistic' terms of a person's beliefs, desires, ideas, hopes, fears, feelings, emotions - is somehow utterly disqualified.

In describing Skinner's suspicion of mentalistic language as "an intuition," Dennett seems to be saying two things:

(1)   that there is a real basis for behaviorist suspicion of our ordinary ways of talking about behaviour, that underlying it there is an important truth, but also

(2)   that behaviorists in general and Skinner in particular have failed to provide any rational justification for their suspicion.

If that is correct, as I believe it is, the only way we can proceed is to have a long and careful look at what contemporary philosophers contemptuously dismiss as "folk psychology" in order to find out how it works and what its 'ontological commitments' are, i.e., what states, events and processes over and above those we can objectively observe are we committed to believing in when we use such language.

Conceptual analysis: its rise and fall

In order to carry out such an investigation of the language of common sense psychology we have to rely, I believe on a technique known as 'conceptual analysis'. Conceptual analysis is a technique for elucidating the meaning of words whose meaning is problematic, which was devised by the Austrian philosopher, Ludwig Wittgenstein, during the period of his later philosophy, the period between 1929, when he finally settled at the University of Cambridge, England, and his death in 1951.

As far as I know, Wittgenstein himself never used the term 'conceptual analysis' and would have been appalled at the suggestion that he was developing something that could be described as 'a technique'. As far as I am aware, the term was first used by the Oxford ordinary language school of philosophy which developed in the immediate aftermath of World War II and whose principal figures were Gilbert Ryle, John Austin, Paul Grice and Sir Peter Strawson, as he now is, the only prominent member of the group who is still alive. Although deeply influenced by the later Wittgenstein, parted as they were by the 77 miles which separate England's two ancient seats of learning, the Oxford group were free to develop their own ideas to a greater extent than those whose proximity put them in thrall to Wittgenstein's dominant personality.

Conceptual analysis, as it developed in the hands of the Oxford group, enjoyed a brief period as the dominant philosophical approach in the English-speaking world. But by the mid-nineteen sixties it was a spent force. This was not due to any failure of the technique as such. It was due rather to two factors.

(1)   the anti-scientific prejudices of the founding fathers of the movement, including Wittgenstein himself, which in the eyes of its critics rendered it unfit to address the brave new intellectual world that was opening up as a consequence of the development of the computer both as a technology and as a model for the functioning of the brain,

(2)   its espousal of the view that all the traditional problems of philosophy are conceptual confusions based on a misunderstanding of how language works. This had what was, for the philosopher, the unappetizing consequence that, once those confusions have been cleared up, all that is left for him to do is to engage in a purely lexicographical/linguistic study of ordinary English for its own sake.

Compared with that, the current conception of the philosopher as the top-down theorist of cognitive science is far more glamorous and far more in keeping with the philosopher's self-image as the universal arbiter of matters intellectual.

This draws our attention to two other respects in which conceptual analysis is in conflict with what, since its demise, has become the dominant view amongst philosophers in the English-speaking world:

(1)   Conceptual analysis is fundamentally at odds with the idea with which Professor Quine has been particularly associated that, to quote my old friend Jack Smart (Professor J. J. C. Smart, personal communication) "the underlying structure of language is that of first order logic."

(2)   Though none of its originators would have conceded this, conceptual analysis is in fact an empirical sociological investigation of the norms or conventions governing the construction and use of the sentences of whatever language or technical code is under investigation.

Philosophers, needless to say, are much happier with the idea that theirs is an *a priori* discipline related to and based on formal mathematical logic, than they are with the idea that it is an empirical discipline, particularly one with affiliations to such a soft and disreputable empirical science as sociology.

Conceptual analysis and the problem of mentalism

But these features that have made conceptual analysis unattractive to the philosopher are features which, if anything, should make it attractive to the behavioral scientist interested in understanding the conceptual tools he is working with or is considering working with. Unfortunately, since its effective abandonment by the philosophers, no one has seen fit to grasp the opportunity which that abandonment presents to develop conceptual analysis as a branch of applied empirical sociolinguistics. Had that happened, the behavioral scientist would now be in a position to take the problem of the meaning of mentalistic terms to the relevant specialist, instead of leaving it to the interminable exchange of insults across the cognitivist/behaviorist divide.

As things stand, all we have to go on are the classic texts of the period when conceptual analysis was being hammered into shape by Wittgenstein and his disciples at Cambridge and by Ryle, Austin, Grice and Strawson at Oxford. Where we are fortunate is in the fact that what Ryle calls "the logical geography of our ordinary mental concepts" is a central pre-occupation of nearly all those classic texts. Where we are less fortunate is in the fact that the most systematic and comprehensive of those texts, Ryle's (1949) book *The Concept of Mind* fails to address the issue of the propriety of using mentalistic language for scientific purposes. This is partly because Ryle was not interested either in science in general or in the development of a science of behavior in particular. But it is also due to the fact that the most obvious conclusion to be drawn from his essentially behaviorist analysis of concepts like 'believing' and 'wanting' is that, once traditional philosophical misunderstandings are cleared out of the way, there is no earthly reason why a behaviorist should not avail him or herself of this subtle and flexible conceptual scheme.

But there is also another reason why Ryle's analysis throws less light on the issue of mentalism than might be hoped. It is this. Ryle's analysis focuses on a particular feature of what as he calls, following Wittgenstein, the "logical grammar" of our ordinary psychological verbs, namely, the feature which linguists call the 'aspect' of the verb. Thus a typically Rylean observation is that while it makes perfectly good sense to talk of someone being engaged for a period of time in looking at, watching, listening to, pondering or calculating something, it does *not* make sense to talk of them being engaged for a period of time in seeing, hearing, recognising, remembering, knowing, believing, wanting or intending. These aspectual phenomena are very important when it comes to working out whether what we talk about when we use a particular psychological verb is an activity, a process, an instantaneous event of the start/stop variety or a dispositional

state. But though they provide a useful background, these distinctions are not the ones that really matter as far as the issue of mentalism is concerned.

Intentionality

By common consent the issue that *is* central to the argument for and against mentalism is one which Ryle does not consider in his book, the issue which has been referred to as 'intentionality' since it was given that name in the last century by Franz Brentano (1874). Intentionality was brought into the focus of discussion within analytic philosophy by Roderick Chisholm (1957) in the nineteen fifties. Since that time, it has been seen as a linguistic phenomenon, a peculiarity of the grammatical object of certain verbs, usually taken to be the verbs of common sense psychology.

Despite the substantial literature which has accumulated on this topic in recent years, the lack of a systematic conceptual analysis of a full range of examples of the kind of sentences which display this phenomenon in all its aspects has meant that we are now confronted more by conceptual confusion on this topic, than by conceptual clarity. Having wrestled with the problem on and off for the past twenty years, I have finally succeeded in publishing a paper (Place 1996) which I hope will finally put the matter to rest.[3] I cannot rehearse all the arguments and evidence here. All I can do is present you with the conclusions I have reached and with that part of the evidence which impinges most directly on the problem of mentalism.

Two kinds of referential anomaly

The general conclusion is that, viewed as a linguistic phenomenon, the term 'intentionality'[4] refers to two distinct kinds of referential anomaly which between them affect the grammatical objects of some, but not all, psychological verbs, all verbs of utterance and all dispositional predicates, i.e., dispositional adjectives, dispositional verbs and the subjunctive forms of non-dispositional verbs (as when we speak of what 'may,' or 'might happen' or 'be done,' or 'might have happened' or 'been done'). A referential anomaly is a linguistic expression which occurs in an argument place within a sentence, typically the direct object position after a transitive verb, which is occupied in the standard case by a noun phrase referring to an actually existing inanimate object or living creature. Such a linguistic expression is anomalous in so far as it does not straightforwardly refer to such an object. The two kinds of referential anomaly involved in what is referred to as intentionality arise in the one case from the use of the expression to *quote* what someone has said or might be expected to say, and in the other case from the use of the expression to indicate the range of events which, if any one of them occurred, would constitute a manifestation or satisfaction of a disposition.

The distinction between these two kinds of referential anomaly is vitally important for the scientific status of mentalistic accounts of behavioral phenomena. That is because dispositions of one sort or another are everywhere in nature from the 'charm' of the quark, through the brittleness of glass and the strength of a response to linguistic competence defined as the ability to construct and construe indefinitely many novel sentences. The referentially anomalous use of language to characterise the range of possible manifestations of such dispositions is the only linguistic device we possess for doing that job. This kind of intentional locution is not merely unobjectionable, it is inescapable.

The quotational variety of referential anomaly is a different matter. As long as the quotation is a quotation of what someone has actually said or actually would say there is no problem. The objectionable case, the one that has stuck in the gullet of behaviorists from the very beginning, is where a quotation or quotational language is used to characterise the behavioral propensity of an organism, an animal, or pre-linguistic human infant, which does not speak, or where it is used to characterise behavior on the part of a linguistically competent adult which is not in fact subject to any kind of linguistic control or "rule-governance," as Skinner (1966) calls it. An example of this kind of misuse of quotational intentionality - or "intensionsality-with-an s", to borrow Searle's (1979; 1983) - is one which I owe to the philosopher Kathy Wilkes (1984, p. 228). It is the example of someone stumbling at the top of a flight of stairs and explaining the stumbling by saying that she thought there was one more step than there actually was. Taken literally the implication is that in order to climb a stair successfully, one needs to construct a verbal rule, specifying the behavior-consequence relation for each step taken.

---

[3] In view of the vested interest of philosophers in perpetuating conceptual confusions which only they can hope to resolve, this hope is almost certainly vain.

[4] Some philosophers prefer to follow the late Professor William Kneale (1968) in using the term "intensionality" when speaking of these features of language, while reserving "intentionality" for the extra-linguistic phenomena which 'intensional' language is used to characterize which is how it was construed by Brentano.

Such locutions, of course, are not intended by the speaker to be taken literally. If I say that my dog thinks he is about to get fed, no one in their right mind would accuse me of believing, with the philosopher Jerry Fodor (1975), that dogs talk to themselves and say to themselves something like "Food is coming" when they anticipate its appearance. Nevertheless, that is the model, the model of what Skinner calls "rule-governed behavior," which is being invoked to make sense of the way the animal behaves.

For the practical purpose of everyday life, interpreting behavior on the implicit assumption that it is controlled by self-directed talk is a harmless metaphor, a convenient and often illuminating fiction. But when the object is to give an accurate scientific account of behavior which invokes only those states, events and processes which actually exist and occur, it is plainly an unacceptable anthropomorphism. And *that* is precisely the charge that the behaviorists have always levelled against the use of mentalistic language.

Sorting the wheat from the chaff

If there are these two forms of intentionality, (a) the dispositional form which is not merely harmless but an indispensible linguistic tool for indicating the scope of a disposition and (b) the quotational form (intensionality-with-an-s) which involves the scientifically unacceptable metaphor of the control of behavior by self-directed talk, when it is applied to the description and explanation of behavior that is not so controlled, it is important for the purposes of the science of behavior that we should be able to distinguish clearly between the two cases. Unfortunately, this is easier said than done. Take for example the sentence

A *Joe wants an apple*

Here it is not difficult to show that the noun phrase *an apple* which occupies the direct object position after the verb is referentially anomalous. All we have to do is compare it with its role in the sentence

B *Joe eats an apple*

In a case where **B** is true *an apple* refers to an actual apple which Joe eats, whereas **A** can be true when there is in fact no actual apple anywhere in the vicinity, and no one particular apple that Joe can be said to want. But there are two possible explanations of this anomaly.

Suppose that *Joe* here is the name of a monkey. In that case the referential anomaly affecting the noun phrase *an apple* in

A *Joe wants an apple*

can be understood in purely dispositional terms. It simply indicates a range of possible objects any one of which, were it to appear, would satisfy Joe's desire and, as *we* would point out, would reinforce the propensity to emit the kind of behavior which immediately preceded that appearance. But if *Joe* is the name of a linguistically competent human being, the description *an apple* can be heard as a quotation, a quotation for example of what Joe actually said when asked what he wanted to eat.

Embedded oratio obliqua *sentences.*

But now consider the case of the sentence

C *Joe believes that he is about to be fed*

Here again, the direct object of the verb is a referentially anomalous expression which describes a range of possible events one of which may or may not actually occur. Here again if *Joe* is the name of a dog this expression simply serves to indicate a range of behaviors the probability of whose emission by the organism in the presence of current discriminative stimuli has been increased by virtue of food having appeared in the past as an immediate consequence of their emission under similar stimulus conditions. And again if *Joe* is a linguistically competent human, it can be taken as a quotation either of something that Joe actually said or, more problematically, as a quotation of what he might have said to himself under his breath.

What creates the problem and makes the difference between this example and

A *Joe wants an apple*

is that the direct object of the verb *believe* in

C *Joe believes that he is about to be fed*

is a complete embedded sentence in *oratio obliqua* or indirect reported speech. In other words, even in the case where Joe is dog, it has the form of an indirect report of a self-directed sentence such as *I am about to be fed* or *Food on its way - Good Oh*! and relies on our intuitive understanding of the kind of behavior that goes with such a self-directed utterance for our ability to anticipate the kind of behavior to be expected of Joe in such circumstances, even in the case where there is no conceivable possibility of such a self-directed utterance having actually occurred.

The conclusion I draw from this is that it is primarily these embedded *oratio obliqua* sentences which we should be suspicious of and, wherever possible, discard, when attempting to give an accurate scientific account of behavior. From this it also follows that we ought to discard verbs like *believe* and *think* which invariably take such an embedded *oratio obliqua* sentences in the grammatical object position, except in the case of the locutions such as *believe in* which has no application to the behavior of pre-linguistic organisms for the very good reason that you can only be said to believe in the existence and/or potency of something, if that existence or potency is something that has been communicated to you by word of mouth.

No such difficulties arise in the case of a verb like *want* which doesn't take an embedded sentence in *oratio obliqua* as its grammatical object, except in the doubtful case of the locution *wanting to do something* which can be interpreted as an embedded *oratio obliqua* sentence in the imperative mood used to quote a self-directed *oratio recta* sentence of the form *Do this!* But it can equally well be interpreted as a straightforward dispositional description used to indicate a range of behaviors which, in Skinner's terms, have a high probability of emission under prevailing circumstances.

I conclude that whereas *believe* is a verb which is too deeply entrenched in the linguistic metaphor to be acceptable in a scientific account of contingency-shaped behavior, the same objections do not apply in the case of the verb *want*. *Want,* I believe is a verb which we need have no qualms about using when talking in a scientific context about behavior that is entirely contingency-shaped, where linguistic control plays no part.

Locutions that escape the metaphor of linguistic control

In addition to the verb *want* which in English (outside the perverted English of philosophers) *never* takes an embedded *oratio obliqua* sentence as its grammatical object (in contrast to *wish* which invariably does so) there are a number of English verbs which take descriptive phrases as well as embedded *oratio obliqua* sentences as their grammatical object. These verbs are shown in **bold type** on Table 1.

But whereas the use of the embedded sentence construction compels an interpretation of the sentence as an oblique quotation of what the agent might be expected to say, and although in some cases a descriptive phrase is likewise to be interpreted as an oblique quotation, as it invariably is when it occurs as the object of a verb of utterance such as *describe*, other locutions of this kind are simply ways of indicating the range of possible satisfaction conditions for the disposition in question. If this is correct, there can be no compelling reason to object to the use of forms such as *Joe expects ____ to happen* (as opposed to *Joe expects that ____ is/will be the case*) when giving a scientific account of behaviour that is not verbally controlled. Similar considerations allow us to accept locutions such as *trying to do ____* even though that form, when it occurs with a verb of utterance as in *Mary asked Joe to do ____*, is the form taken by an oblique quotation of an imperative.

Conclusion

In our account of associative learning theory we saw that in order to state the Rescorla-Wagner model for classical/respondent conditioning, we need to invoke the principle whereby if the onset of a stimulus S_1 (the CS) regularly precedes the onset of another stimulus S_2 (the UCS) an organism will learn when presented with S_1 to *expect, anticipate,* or *predict,* the onset of S_2. We also saw that in order to account for the results of the reinforcer-devaluation experiment, we are confronted with a choice between Tony Dickinson's theory which holds that what the animal *learns* in an instrumental/operant learning situation is to *believe that* pressing the lever will result in the delivery of a particular foodstuff and that what it *does* depends on whether it *wants* or *does not want* that foodstuff, and Bob Rescorla's theory which holds that what the animal learns is the same as that in classical/respondent conditioning, namely to *expect, anticipate,* or *predict,* a succession of stimulus events in the order of the three term-contingency *stimulus - response* (or rather, stimulus feedback from the emission of the response) and *outcome* or *consequence*. What we can now see from an examination of Table 1 is that, of the psychological verbs thrown up by the Rescorla-Wagner analysis of classical/respondent conditioning, two, *expect* and *anticipate*, are printed in bold type which means that, when they are used with a simple noun phrase as their grammatical object, in other words when what someone expects or anticipates is an event rather than that a sentence describing it be true, such locutions behave in such a way as to escape the metaphor of linguistic control which are otherwise all pervasive in our ordinary psychological language.

The third verb, the verb *predict* does not break loose from the metaphor, but not because of what Ryle would call its "logical behaviour" which is, I would say, impeccable. The trouble, needless to say, is with

Embedded sentences in AFTER VERBS OF	Indicative (*that...*) Interrogative (*who, why, how,* etc...) Imperative (*to...*)	Indicative (*that...*) & Interrogative (*who, why, how,* etc...)	Indicative (*that...*) & Imperative (*to...*)	Interrogative (*who, why, how,* etc...) & Imperative (*to...*)	Indicative (*that...*) only	Interrogative (*who, why, how,* etc...) only	Imperative (*to...*) only
UTTERANCE	tell advise	say suggest explain	admit swear promise agree mean allow	ask	inform answer assert deny declare submit assure demand concede reply repeat remark add insist	question (*whether...*) describe (*how...*)	request order command undertake authorise permit
COGNITION 1.	**learn know remember forget**	**see hear notice** predict	**expect be reminded** ----------[of] **think** ---[of, about]		**recognise anticipate imagine** believe ------------[in]		
COGNITION 2.		guess wonder predict estimate work out calculate doubt (*whether...*)			dream (tr.) surmise suppose infer conclude		
VOLITION 1.			**be pleased** ----------[at] **be afraid** ----------[of] **be sad** --[dened by]		**be excited** ----------[by] **be angry** ----[with, at] **be worried** ----------[by] **be disgusted** ----------[at] **be unhappy** ----------[at] **be relieved** ----------[at]		**like want**
VOLITION 2.	decide		wish intend				**try**

Table 1. English verbs classified according to the mood of an embedded sentence which can occur as its grammatical object. Psychological verbs which also accept as grammatical object a noun phrase describing an object or event are classified either as COGNITION 1. or as VOLITION 1. Those which occur (transitively) only with an embedded sentence in the object position are classified COGNITION 2. or VOLITION 2. Those usages of psychological verbs which appear to escape the metaphor of linguistic control are given in **bold type**

its etymology deriving as it does from the Latin verb *dicere* - to say.

Turning to the dispute between Dickinson and Rescorla over the implications of the reinforcer devaluation experiment for the analysis of instrumental/operant learning, we see that Dickinson's proposal which is to use the concepts of *believing that*, and *wanting*, while escaping criticism as far as its use of the concept of *wanting* is concerned, involves an unacceptable introduction of the notion that the behavior in question is controlled by a verbal formula in its use of the verb *believe*, particularly in the form of *believe that* which invariably introduces an embedded *oratio obliqua* sentence in the indicative mood.

On the other hand, Rescorla's proposal which combines the Rescorla-Wagner analysis of classical/respondent conditioning with the three-term contingency, requires no new mentalistic features for its formulation over and above those already mentioned under Rescorla and Wagner, the verbs *expect*, *anticipate,* and *predict*, of which only the latter is questionable. It is arguable, of course, that in order to account for instrumental/operant *behavior*, as opposed to instrumental/operant *learning*, we need to talk about whether the organism *likes* or *dislikes, wants* or *does not want* the outcome or consequence of the behavior in question. But as we have already seen, and as is indicated by the fact that they are printed in bold type on Table 1, *like* and *want* are among the locutions that escape the linguistic metaphor.

However, in order to reconcile the hard-nosed radical behaviorist to the use of verbs such as *expect*, *anticipate, like* and *want* in the analysis of behavior, it is not enough to show, as I hope I have, that the use of such terms does not bring with it, as does the use of most mentalistic language, the implication that the behavior in question is, in Skinner's terms "rule-governed", subject, in other words, to stimulus control by a contingency-specifying verbal formula.

What the hard-nosed radical behaviorist will need to know before he or she can accept the use of such terms with equanimity is what kind of private events are we attributing to the organism when we use these words. Skinner always insisted during his lifetime that radical behaviorism recognises both the existence of private events and the possibility of studying them scientifically through the verbal reports of their human owners. What he has not left us is much in the way of guidance as to their nature and function within the behavioral process. It is my conviction that, in the face of the kind of experimental evidence which has been accumulated over recent years by those working within the framework of associative learning theory, that is a gap which needs urgently to be filled. But how to fill it?

Well, if the primary evidence for the existence and nature of private events comes from the verbal reports of human subjects, and if, as we surely must, we reject the reports of psychologists and behavioral scientists, of whatever persuasion, on the grounds that their reports will inevitably be contaminated by their theoretical pre-conceptions, we have to content ourselves with the verbal reports of naive subjects. Inevitably those reports will be formulated, not in the language of behavioral science, but in the language of common sense psychology. Consequently, in order to evaluate those reports, we are going to need professional help, the professional help of the conceptual analyst.

As a conceptual analyst who has been wrestling with this problem for more than forty years, I have some ideas about the kinds of private event which such an investigation is likely to reveal, but that's another story.

References

Adams, C. D. and Dickinson, A. (1981) Instrumental responding following reinforcer devaluation. *Quarterly Journal of Experimental Psychology*, **33 B**: 109-112.

Brentano, F. (1874). *Psychologie vom empirischen Standpunkt*. Duncker & Humblot.

Brogden, W. J. (1939) Sensory preconditioning. *Journal of Experimental Psychology*, **25**, 323-332.

Brown, P. L. and Jenkins, H. M. (1968) Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior*, **11**, 1-8.

Chase, S. (1968) Selectivity in multidimensional stimulus control. *Journal of Comparative and Physiological Psychology*, **66,** 787-792.

Chisholm, R.M. (1957) *Perceiving: A Philosophical Study*. Ithaca, NY: Cornell University Press.

Dennett, D.C. (1978) *Brainstorms: Philosophical Essays on the Mind and Psychology*. Montgomery, Vermont: Bradford.

Dickinson, A. (1988) Intentionality in animal conditioning. In L. Weiskrantz (ed.) *Thought without Language*. Oxford: Clarendon Press, pp. 305-325.

Dickinson, A. and Balleine, B. (1993) Actions and responses: The dual psychology of behaviour. In N. Eilan, R. A. McCarthy, & B. Brewer (Eds.), *Spatial representation: Problems in philosophy and psychology* (pp. 277–293). Blackwell Publishing.

Edelman, G.M. (1987) *Neural Darwinism: The Theory of Neuronal Group Selection*. New York: Basic Books.

Festinger, L. (1957) *A theory of cognitive dissonance.* Stanford University Press.

Fodor, J. (1975) *The Language of Thought.* New York: Crowell.

Freeman, F. and Thomas, D.R. (1967) Attention vs. cue utilization in generalization testing. Paper presented at Midwestern Psychological Association, Chicago.

Heyes, C. and Dickinson, A. (1990) The intentionalty of animal action. *Mind and Language, 5,* 87-104.

Hilgard, E. R. and Marquis, D.G. (1940) *Conditioning and Learning.* New York: Appleton-Century-Crofts.

Kneale, W. (1968) Intentionality and intensionality. *Proceedings of the Aristotelian Society,* Supplementary Volume XLII, 73-90.

MacPhail, E. M. and Reilly, S. (1989) Rapid acquisition of a novelty versus familiarity concept by pigeons (*Columba livia*). *Journal of Experimental Psychology (Animal Behaviour Processes), 15,* 242-252.

Morgan, C. Lloyd (1894) *Introduction to Comparative Psychology.* London: Scott.

Pavlov, I. P. (1927) *Conditioned Reflexes.* English Translation by G. V. Anrep. London: Oxford. University Press.

Place, U. T. (1954) The concept of heed. *British Journal of Psychology 45,* 234-55. Reprinted in D. F. Gustafson, (ed.), *Essays in Philosophical Psychology.* New York: Doubleday, 1964, pp.206-226.

Place, U. T. (1985) A response to Sundberg and Michael. *The Analysis of Verbal Behavior 3,* 38-45

Place, U. T. (1992a) Eliminative connectionism and its implications for a return to an empiricist/behaviorist linguistics. *Behavior and Philosophy, 20,* 21-35.

Place, U. T. (1992b) 'Selectionism and connectionism: their implications for a return to an empiricist/behaviorist linguistics.' Paper presented at the 1992 Conference of the Sociolinguistics Research Committee for Sociolinguistics of the International Sociological Association on 'The interface between Sociology and Linguistics', Katholieke Universiteit Nijmegen, 9th-11th June 1992.

Place, U. T. (1996) 'Intentionality as the mark of the dispositional' *Dialectica, 50,* 91-120.

Rescorla, R. A. (1991) Associative relations in instrumental learning: The eighteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology, 43B:* 1-23.

Rescorla, R. A. and Wagner, A. R. (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black and W. F. Prokasy (eds.) *Classical Conditioning,* Vol. 2: *Current Research and Theory.* Englewood Cliffs, NJ: Prentice-Hall.

Ryle, G. (1949) *The Concept of Mind.* London: Hutchinson.

Searle, J. R. (1979) 'What is an Intentional State?' *Mind,* LXXXVIII, 74-92.

Searle, J. R. (1983) *Intentionality: an Essay in the Philosophy of Mind.* Cambridge: Cambridge University Press.

Skinner, B. F. (1938) *The Behavior of Organisms.* New York: Appleton-Century.

Skinner, B. F. (1957) *Verbal Behavior.* New York: Appleton-Century-Crofts.

Skinner, B. F. (1966) An operant analysis of problem solving. In B. Kleinmuntz (Ed.) *Problem Solving: Research, Method and Theory,* Wiley. Reprinted as Chapter 6 of Skinner, B.F. (1969). *Contingencies of Reinforcement: A Theoretical Analysis.* Appleton-Century-Crofts.

Skinner, B. F. (1989) The behavior of the listener. In S. C. Hayes (Ed.) *Rule-Governed Behavior: Cognition, Contingencies and Instructional Control.* New York: Plenum, Chapter 2, pp. 85-96.

Tolman, E. C. and Honzik, C. H. (1930) Introduction and removal of reward, and maze performance in rats. *University of California Publications in Psychology, 4,* 257-275.

Wilkes, K. (1984) Is consciousness important? *British Journal of the Philosophy of Science, 35,* 223-243.

Wittgenstein, L. (1953) *Philosophische Untersuchungen.* English translation as *Philosophical Investigations* by G. E. M. Anscombe. Oxford: Blackwell.

Wittgenstein, L. (1958) *The Blue and Brown Books.* Oxford: Blackwell.